Neuron
Article

Neurons in the Ventral Striatum Exhibit Cell-Type-Specific Representations of Outcome during Learning

Hisham E. Atallah,[1] Andrew D. McCool,[1] Mark W. Howe,[1] and Ann M. Graybiel[1,*]
[1]McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
*Correspondence: graybiel@mit.edu
http://dx.doi.org/10.1016/j.neuron.2014.04.021

SUMMARY

The ventromedial striatum (VMS) is a node in circuits underpinning both affect and reinforcement learning. The cellular bases of these functions and especially their potential linkages have been unclear. VMS cholinergic interneurons, however, have been singled out as being related both to affect and to reinforcement-based conditioning, raising the possibility that unique aspects of their signaling could account for these functions. Here we show that VMS tonically active neurons (TANs), including putative cholinergic interneurons, generate unique bidirectional outcome responses during reward-based learning, reporting both positive (reward) and negative (reward omission) outcomes when behavioral change is prompted by switches in reinforcement contingencies. VMS output neurons (SPNs), by contrast, are nearly insensitive to switches in reinforcement contingencies, gradually losing outcome signaling while maintaining responses at trial initiation and goal approach. Thus, TANs and SPNs in the VMS provide distinct signals optimized for different aspects of the learning process.

INTRODUCTION

The brain circuits that process reinforcement exert powerful influences over behavior: they enable normal responsiveness to rewarding and aversive stimuli, mediate reinforcement-based learning, and have been implicated as dysfunctional in neuropsychiatric disorders ranging from depression to addictive states (Alexander et al., 2010; Belin et al., 2013; Der-Avakian and Markou, 2012; Greening et al., 2013). Findings in humans and experimental animals have highlighted the striatum as a nodal part of this reinforcement-related circuitry important for learning (Brown et al., 2012; Kreitzer and Malenka, 2008; Warner-Schmidt et al., 2012). This view is consistent with the position of the striatum as a principal recipient of signals from midbrain dopamine-containing neurons that could mediate reinforcement-based learning (Dezfouli and Balleine, 2012; Gläscher et al., 2010; Graybiel, 2008; Hyman et al., 2006; Liljeholm and O'Doherty, 2012; Tobler et al., 2003; Wickens et al., 2007; Yin and Knowlton, 2006). The ventromedial striatum (VMS), in particular, is considered key to the early stages of reinforcement-based learning (Belin and Everitt, 2008; Belin et al., 2009; Graybiel, 2008; Morris et al., 2004; Pan et al., 2005; Roesch et al., 2007; Schultz, 2002; van der Meer et al., 2010), and neural signals in the VMS are sensitive both to the actions resulting in rewarded outcomes as well as to the outcomes themselves (Li and Daw, 2011; McClure et al., 2003). In rodents, lesions and inactivations of the VMS impair the capacity of the animals to learn how to obtain reward, especially early in training (Atallah et al., 2007; Chang et al., 2012; Everitt and Robbins, 2005; Gill et al., 2010; Hernandez et al., 2002).

In parallel with its key function in the early stages of reinforcement learning, the VMS is centrally implicated in regulating mood and affect and in the etiology of neuropsychiatric conditions (Berridge et al., 2010; Brown et al., 2012; Der-Avakian and Markou, 2012; Greening et al., 2013; Hyman et al., 2006; Keedwell et al., 2005; Kreitzer and Malenka, 2008; Nestler and Carlezon, 2006; Price and Drevets, 2010; Tan et al., 2012; Wacker et al., 2009; Warner-Schmidt et al., 2012). These mood- and motivation-related functions of the VMS probably reflect its interconnections with the dopamine-containing ventral tegmental complex and related nuclei of the brainstem and with limbic forebrain circuits spanning orbitofrontal cortex, medial prefrontal cortex, hippocampus, and amygdala (Humphries and Prescott, 2010; Voorn et al., 2004).

A remarkable set of findings now suggests that these limbic circuits can be divided into pathways that are specialized for processing either rewarding or aversive outcomes and that they may be implicated differentially in affective disorders (Bromberg-Martin et al., 2010; Brown et al., 2012; Cohen et al., 2012; Kreitzer and Malenka, 2008; Setlow et al., 2003; Shabel et al., 2012; Stamatakis and Stuber, 2012; Tan et al., 2012; Warner-Schmidt et al., 2012). Within the VMS, cholinergic interneurons have been found to be critical for the control of affective states by virtue of their selective expression of the serotonin receptor-interacting protein, p11 (Warner-Schmidt et al., 2012). Deletion of p11 or silencing of the cholinergic interneurons precipitates anhedonia-like states in mouse models, and p11 expression is diminished in human depression (Alexander et al., 2010; Svenningsson et al., 2006; Warner-Schmidt et al., 2012). Moreover, these cholinergic interneurons are the specific

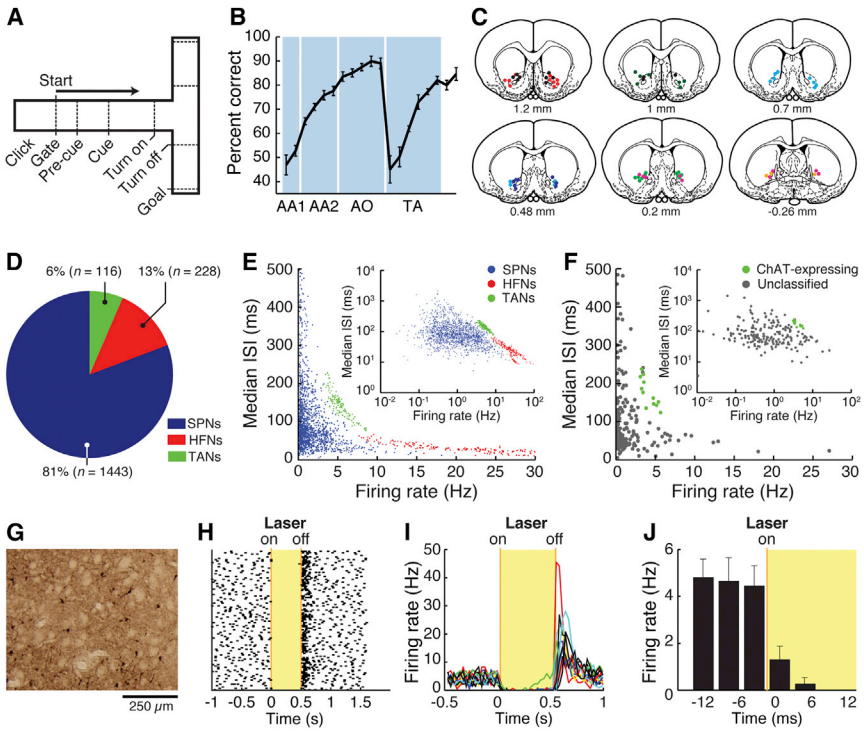

Figure 1. Behavioral Task and Classification of Recorded Neuronal Populations

(A) The T-maze task.

(B) Average learning curve for all animals (n = 8) across the four successive training phases: early (AA1) and late (AA2) auditory task acquisition, auditory task overtraining (AO), and tactile task acquisition (TA). Error bars represent SEM.

(C) Tetrode recording sites. Colors represent sites for individual rats. Anterior-posterior coordinates indicated below.

(D) Percentage of recorded VMS units classified as spiny projection neurons (SPNs, blue), high-firing neurons (HFNs, red), and tonically active neurons (TANs, green).

(E) Firing rates and median ISIs for recorded units, color coded by putative cell type, distinguishing three clusters occupying distinct domains. Axes truncated to show regions of overlap. Insert: axes in log scale to show results for all recorded units.

(F) Firing rates and median ISIs of optogenetically identified cholinergic interneurons (n = 12, green) and unclassified units (gray) in ChAT-Cre transgenic mice (n = 6), plotted as in (E).

(G) Photomicrograph illustrating striatum of ChAT-Cre mouse stained with anti-GFP antibody showing cell bodies (dark brown) and processes (lighter brown) of transfected cholinergic interneurons.

(H) Spike raster plot of a cholinergic interneuron silenced by a 500 ms yellow light pulse (yellow shading).

(I and J) Spike suppression in 12 individual cholinergic interneurons (I) and average suppression (J) that occurred rapidly in response to the light pulse. Error bars represent SEM.

See also Figure S1.

targets of GABA-containing neurons of the ventral tegmental area (VTA) that promote aversive responses, and optogenetic stimulation of this pathway enhances discrimination of aversive conditioning stimuli from neutral stimuli (Brown et al., 2012; Tan et al., 2012). Here we identify unique dynamics of valence signaling by populations of VMS tonically active neurons (TANs) that have firing properties resembling those of optogenetically identified cholinergic interneurons, and we contrast these with the strikingly different learning-related firing properties of VMS spiny projection neurons (SPNs) and high-firing interneurons (HFNs). The expression of these subtype-specific signals could be critical to a mechanistic account of how VMS circuits participate both in reinforcement learning and in goal-directed behavior and in the control of affective states.

RESULTS

We recorded the single-unit activity of neurons identified as putative TANs, SPNs, and HFNs with multiple tetrodes placed bilaterally in the VMS of rats that consecutively learned two versions of an instrumental T-maze task. At the switch between the first and second versions, the modality of the cues instructing the location of reward, representing the reinforcement contingency of the task, was changed from auditory or tactile. The maze context and actions required were otherwise identical (Barnes et al., 2005; Kubota et al., 2009; Thorn et al., 2010). With this design, we asked whether the firing of the VMS neurons tracked changes in trial-by-trial outcomes that we observed during learning of the first task version and then during learning of the second task version after the reinforcement contingency switch, or whether they tracked generalized learning of goal-directed behavior within the maze context, which could be accomplished by acquisition of the first version alone.

The behavioral task and recording arrangements are shown in Figures 1A and 1C (see also Experimental Procedures and Supplemental Information available online). The rats were trained initially with auditory instruction cues in daily ca. 40-trial sessions until they reached an acquisition criterion of three out of four consecutive sessions above 72.5% correct performance. The first and second acquisition phases (AA1 and AA2) started, respectively, on the first day of training and when the rat reached 60% correct performance (Figure 1B; Experimental Procedures). The animals were then overtrained for an additional ten sessions on the auditory task (phase AO). For six rats, training was then continued without interruption on the second version of the task, in which tactile cues replaced the auditory cues, and the rats were again trained to the same learning criterion (phase TA; Figure 1B; Supplemental Experimental Procedures). The rats reached the learning criterion on the auditory version of the task in 8–23 sessions (mean = 15.5) and learned the subsequent tactile version of the task in 8–13 sessions (mean = 10.7). The rate of change in performance on the task was significantly greater during the training on the second tactile task version (Mann-Whitney Wilcoxon rank-sum test, $p < 0.05$).

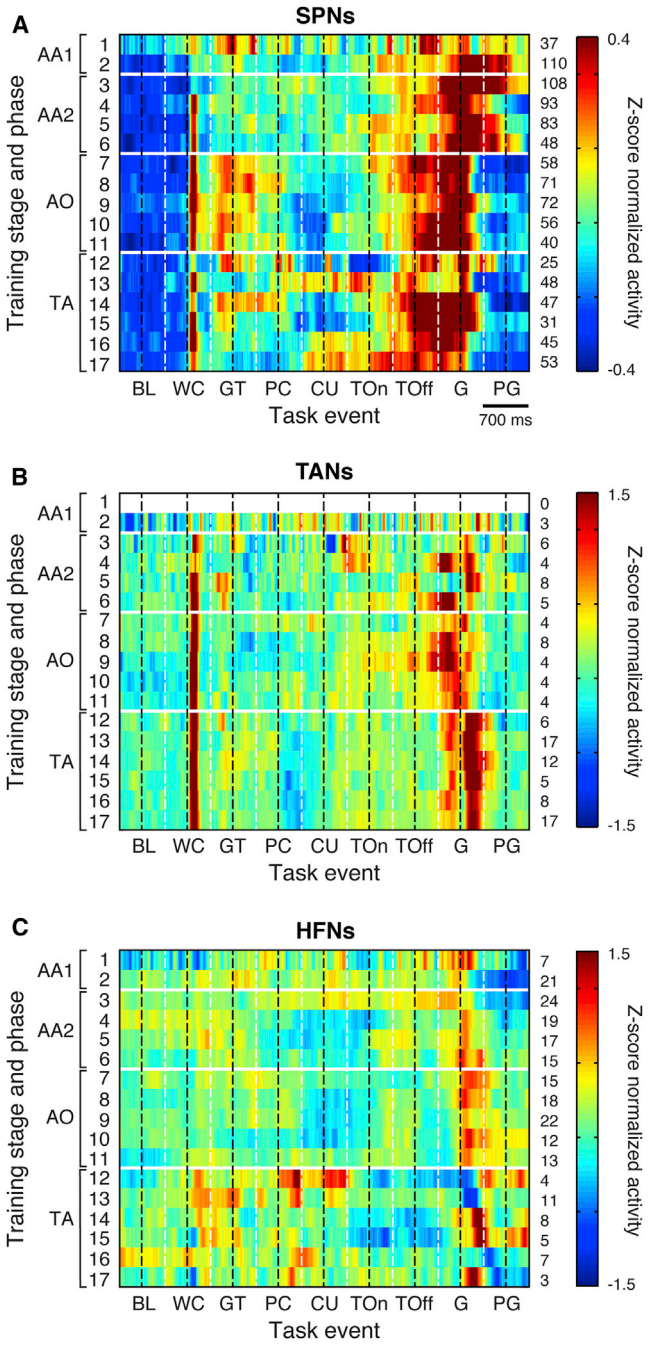

Figure 2. Ensemble Activity of SPNs, TANs, and HFNs during Training

(A–C) Pseudocolor plots showing average *Z* score normalized firing rates on rewarded trials for task-responsive SPNs (A), all recorded TANs (B), and all recorded HFNs (C), plotted in 20 ms bins, relative to baseline firing rates. Successive ±350 ms perievent windows (BL, baseline; WC, warning click; GT, gate opening; PC, precue; CU, cue onset; TOn, turn onset; TOff, turn offset; G, goal reaching; PG, postgoal) were abutted in order of occurrence within a trial; continuous time is not shown. Training stages and phases (indicated to left) as defined in text. Number of units compiled for each stage indicated at right. See also Figure S2.

Running times decreased across training on the auditory version of the task (Kruskal-Wallis ANOVA; $p < 0.01$), and they increased for only a single session when training with tactile cues began (Figure S1A; Dunn-Sidak post hoc test, $p < 0.05$).

Of the 1,787 well-isolated units accepted for analysis, 81% were identified as putative SPNs (Figure 1D; Supplemental Experimental Procedures). Among the remainder, we identified putative HFNs, accounting for 13%, and putative TANs, accounting for 6%. These populations were distinguished according to conventional criteria and on the basis of their different relationships between mean firing rate and median interspike interval (ISI): even when the mean firing rates of the groups overlapped, their median ISIs differentiated them (Figure 1E). To rule out the possibility that task responses were contributing to the separation between cell types, we replicated this finding using spikes collected in a 2 s baseline period before each trial (Figure S1B).

In a further effort to identify the TAN population, we compared the mean firing rates and median ISIs of the TANs that we recorded in these rats to those of cholinergic interneurons that we identified optogenetically in six transgenic ChAT-Cre mice (n = 12 units, Figure 1F) (Gong et al., 2007) in which a Cre-dependent AAV5 virus had been injected into the striatum (Figure 1G; Supplemental Experimental Procedures). Units that were suppressed by a 500 ms pulse of yellow light were considered to be cholinergic interneurons (Figures 1H–1J) (English et al., 2012). These optogenetically identified interneurons fired in the 3–6 Hz range and displayed ISIs distinguishing them from the nonlight suppressed population, presumably the SPNs and GABAergic interneurons (Figure 1F). The firing rates and ISI distributions for putative TANs in the rats were equivalent to those in the mice. All findings refer to these identifications of putative neuronal classes, with the acknowledgment that other cell types could have been encountered in our sample.

VMS Ensemble Activity Marks Beginning and End Periods of the Maze Runs and Subpopulations of VMS Neurons Exhibit Ramping Activity as Goal Is Approached

All three classes of VMS neuron that we identified exhibited task-related activity observable in the net activity of each class, illustrated in Figure 2 as perievent time histograms (PETHs) for correctly performed trials collapsed across training stages. PETHs were constructed for 700 ms windows centered on each task event and thus do not represent uninterrupted time between task events. The discontinuities in task time were prevalent early in training when trial durations were relatively long (Figure S1A).

Both the task-responsive SPNs, which consisted of 72% of the total SPN population (Figures S2A and S2B; Supplemental Experimental Procedures), and the TAN ensembles (Figure 2B) exhibited a sharp peak in response to the warning click signaling trial start and intense activity at the time of goal reaching. SPN firing exhibited a gradual increase in activity during the maze run (Figure 2A). Neither of these responses was observed in the non-task-responsive SPNs (Figures S2C and S2D). Activity in between these times occurred mainly around gate opening, as the runs started, and at the approach to the goal. In contrast to the SPNs and TANs, the HFN ensembles exhibited little activity at the warning click. HFNs became active near the beginning of each run and again at goal reaching, when the maze run was

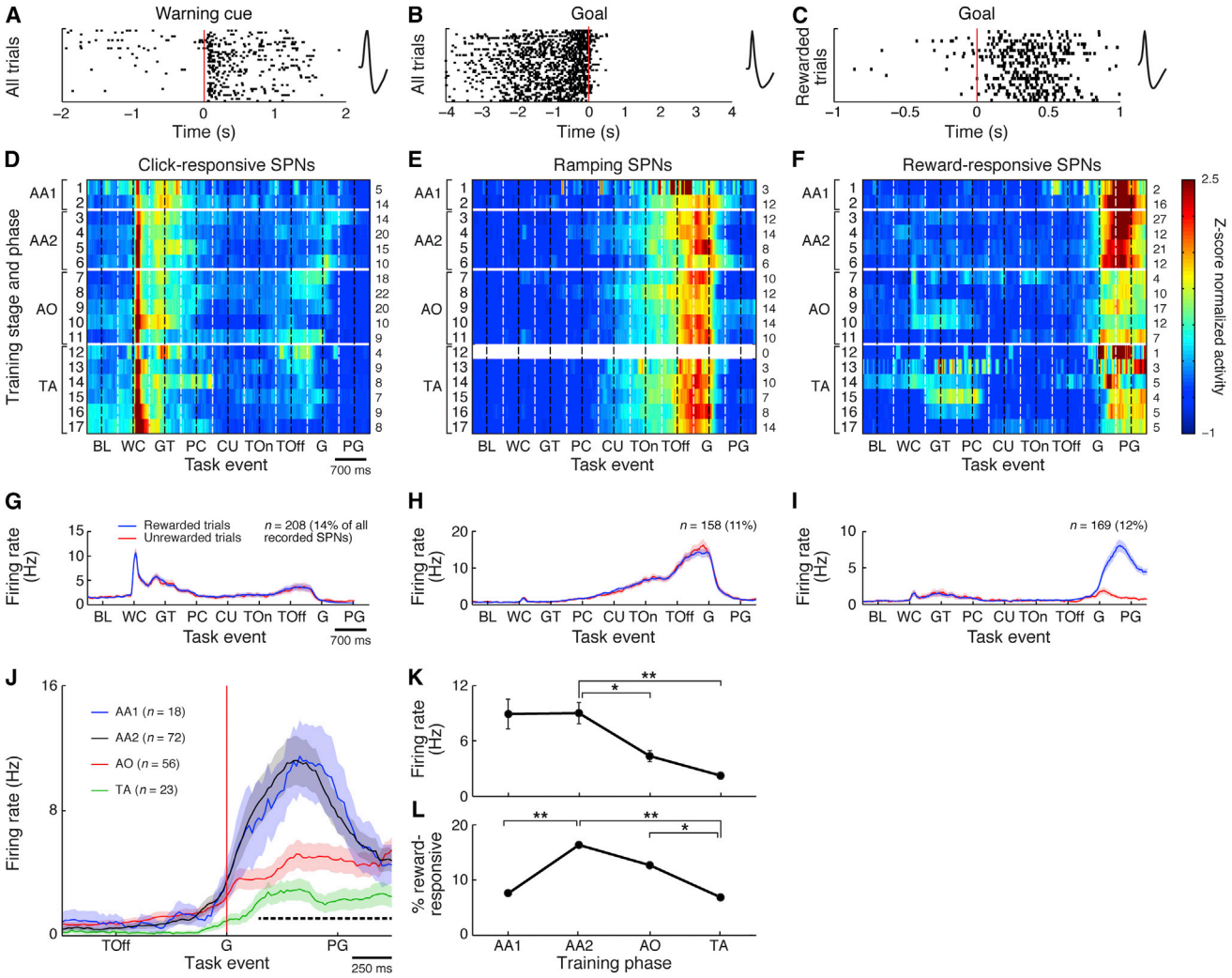

Figure 3. Three Types of SPN Response Recorded during Training

(A–C) Raster plots for click-responsive (A), ramping (B), and reward-responsive (C) SPNs. The average waveform for each unit is shown at right.

(D–F) Average *Z* score normalized activity of click-responsive (D), ramping (E), and reward-responsive (F) SPNs across training stages, plotted as in Figure 2.

(G–I) Firing rates of click-responsive (G), ramping (H), and reward-responsive (I) SPNs in rewarded (blue) and unrewarded (red) trials, averaged across all training stages. Shading indicates SEM.

(J) Firing rates of reward-responsive SPNs in rewarded trials, averaged separately for each training phase (color coded). Dashed line represents the period of licking behavior.

(K) Average firing rate of reward-responsive SPNs during a 700 ms period after goal reaching across training phases. *$p < 0.05$; **$p < 0.01$. Error bars represent SEM.

(L) The percentage of reward-responsive SPNs across training phases.

See also Figures S3 and S4.

complete and reward delivery occurred on correctly performed trials (Figure 2C). Thus, the ensemble activities of all three classes of VMS neurons emphasized mainly the early and late periods of the maze runs, a characteristic previously found for dorsal striatal neurons, especially in the dorsolateral striatum (DLS) (Barnes et al., 2005; Jog et al., 1999; Thorn et al., 2010).

VMS Spiny Projection Neurons Exhibit Three Main Response Profiles

We first identified subsets of task-responsive SPNs that had selective responses to the warning click (click-responsive SPNs, n = 208, 14% of all recorded SPNs), during approach toward the goal location (ramping SPNs, n = 158, 11%), or during the reward period (reward-responsive SPNs, n = 169, 12%; Figures 3A–3I; Supplemental Experimental Procedures). Click-responsive SPNs were selected based on their increased firing rate during a 240 ms window after the click, relative to their baseline firing (Figures 3D and 3G). SPNs were classified as having ramping responses if they increased their firing in four successive windows (turn off > turn on > cue > precue) (Figures 3E and 3H). We identified as reward-responsive SPNs those that were phasically activated in the goal-reaching period and that

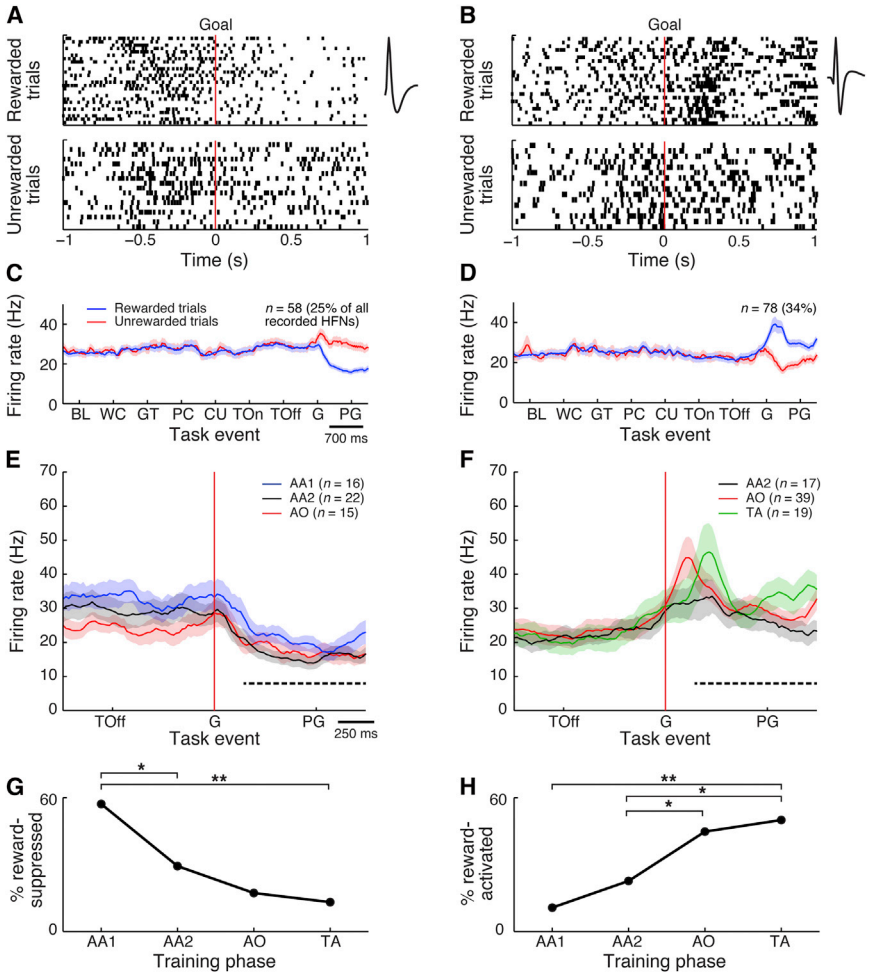

Figure 4. Reward-Responsive High-Firing Neurons

(A and B) Spike activity of a reward-suppressed HFN (A) and a reward-activated HFN (B). Their average waveforms are shown at right.

(C and D) Firing rates of reward-suppressed (C) and reward-activated (D) HFNs in rewarded (blue) and unrewarded (red) trials, averaged across all training stages. Shading indicates SEM.

(E and F) Average firing rates of reward-suppressed (E) and reward-activated (F) HFNs in rewarded trials, plotted as in Figure 3J. Data from training phases with fewer than five units are not shown.

(G and H) The percentage of reward-suppressed (G) and reward-activated (H) HFNs across training phases. *p < 0.05; **p < 0.01.

See also Figure S4.

fired significantly more for correct trials than for incorrect trials during a 700 ms window after goal reaching (Figures 3C and 3F; Supplemental Experimental Procedures). Despite their definition based on responses to click and goal events, a majority of the click-responsive (81%) and the reward-responsive (88%) SPNs responded to more than one task event. Ramping SPNs had a higher proportion of multievent responses (94%), given their steady increases in firing rate during the maze runs. The spiking of units in all three subsets of VMS neurons exhibited significant correlations with running speed (73%) and position on the maze (82%) when compared with the firing rates in a shuffled data set, but these correlations were extremely weak (Figures S3A and S3B; 95% of units with a correlation coefficient $R^2 < 0.08$). A subset of units (17%) had significantly different firing rates during right and left turns (Figures S3C and S3D). These neurons included 55% of the neurons classified as ramping SPNs.

Reward-Related Activity of VMS Spiny Projection Neurons Declined Steadily across Training

During initial training, reward-responsive SPNs increased their firing rates phasically before and during the occurrence of licking behavior in rewarded trials (Figures 3I and 3J), but this phasic activity nearly disappeared during auditory overtraining, despite the continuation of reward delivery and licking responses after correct performance (Figures 3F and 3J). We averaged the raw firing rates of these reward-responsive SPNs for the 700 ms window after goal reaching for each training phase (Figure 3K). In contrast to the SPNs with click or ramping responses, the reward-responsive SPNs reduced their activity during the goal window in the overtraining from high initial levels, and their responses continued to decline, nearly monotonically, after the cue switch (Figure 3K; Kruskal-Wallis ANOVA; $p < 0.05$). The proportions of reward-responsive SPNs among the total population of SPNs also fell, and they remained low after the initial training phase (Figure 3L; chi-square test; $p < 0.05$). Thus, as training progressed, the population response of the SPNs at the time of reinforcement fell and did not recover during subsequent acquisition of the tactile version of the task.

Subpopulations of HFNs in the VMS Exhibited Opposing Responses at Outcome and Were Oppositely Modulated during Training

To examine possible local circuit activity patterns that could account for the loss of outcome signaling by the reward-responsive SPNs, we examined the outcome period activity of the HFNs. By virtue of their high firing rates, these were candidate inhibitory interneurons (Kawaguchi et al., 1995; Kreitzer, 2009). Over half of the HFNs fell into one of two distinct subgroups whose response properties at reward were of opposite polarity during the 700 ms time window after goal reaching: 25% of the recorded HFNs (n = 58) fired significantly less in rewarded trials than in unrewarded trials (Figures 4A, 4C, 4E, and 4G), whereas 34% (n = 78) fired significantly more during rewarded trials than unrewarded trials (Figures 4B, 4D, 4F, and 4H). Neither of these two HFN populations was responsive to the warning click (Figures 4C and 4D). Many (43%) exhibited different response levels

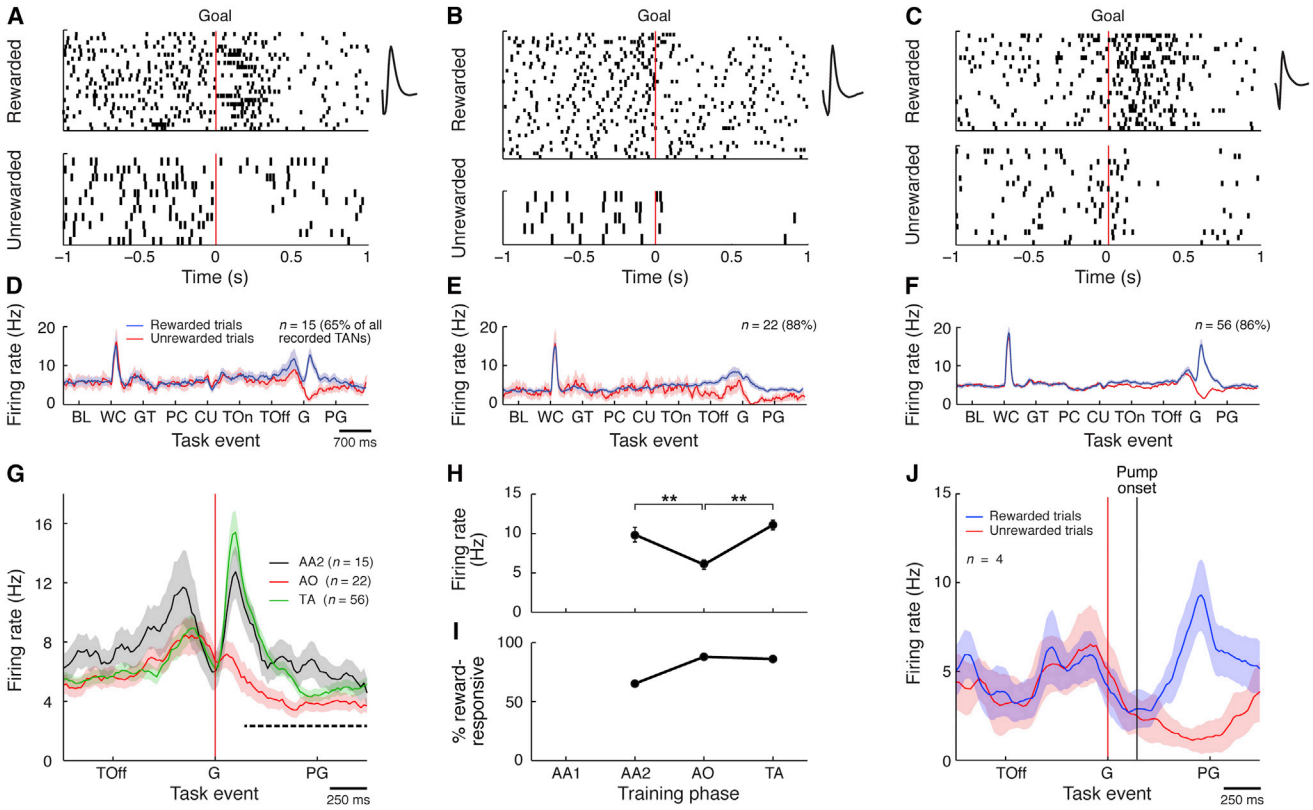

Figure 5. Reward-Responsive Tonically Active Neurons

(A–C) Spike activity of three individual reward-responsive TANs recorded in training phases AA2 (A), AO (B), and TA (C) around goal reaching in either rewarded (top) or unrewarded (bottom) trials. An example from the initial acquisition phase AA1 is not shown because only three TANs were recorded in this phase. The average waveform of each unit is shown on the right.

(D–F) Average firing rate of all reward-responsive TANs in training phases AA2 (D), AO (E), and TA (F) in rewarded (blue) and unrewarded (red) trials. Shading indicates SEM.

(G) Average firing rate of reward-responsive TANs in rewarded trials across training phases, plotted as in Figure 3J.

(H) The average firing rate of the reward-responsive TANs on rewarded trials measured in a 300 ms period after goal reaching across training phase. Error bars show SEM. **$p < 0.01$.

(I) Percentage of reward-responsive TANs across training phases.

(J) Average firing rate of reward-responsive TANs in response to delayed reward (pump onset).

See also Figure S4.

for left and right turns, but the magnitude of the difference in their responses was very weak (Figure S3E). The two populations exhibited a similar relationship between their mean firing rates and median ISIs (Figure S4A).

Over the course of training, the relative proportions of these subpopulations shifted. The proportions of HFNs that exhibited suppressed activity during reward decreased from 57% during initial acquisition to 13% in the tactile training phase (Figure 4G; chi-square test, $p < 0.001$). Conversely, the HFNs that increased their firing rates during reward increased from 10% during initial acquisition to over 50% during the tactile task (Figure 4H; chi-square test, $p < 0.001$), and they exhibited a small but statistically significant delay in their responses during tactile acquisition (Figure 4F). There were almost no changes in firing rate for the reward-suppressed HFNs (Figure 4E). The firing rates of the reward-activated HFNs increased during overtraining, and they maintained this activity during tactile acquisition.

TAN Responses in the VMS Reported Reinforcement Outcome Selectively during Acquisition Phases of Training

In contrast to the nearly monotonic changes in signaling across training of the reward-responsive SPNs and HFNs, the reward-related responses of the TANs recorded exhibited striking modulation in their response patterns during the successive training phases, and these responses changed sharply at task switch (Figures 2B and 5). The reward-related responses of the TANs during the 300 ms window after goal reaching were prominent during training on the auditory task, then fell during overtraining, and then rose again after the switch to the task version with tactile instruction cues. Moreover, the reward-responsive TANs were unique in exhibiting bidirectional outcome signaling: they increased their firing rates for rewarded trials and decreased their firing rates for nonrewarded trials (Figures 5A–5F). These TANs exhibited the most homogeneous response

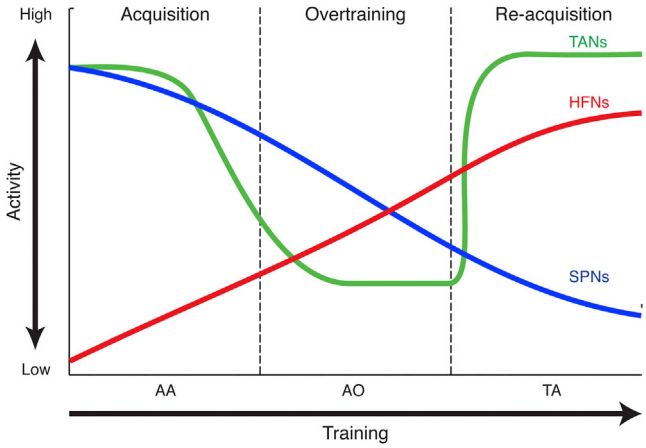

Figure 6. Changes in Activity of Different Neuronal Subtypes during Learning

Schematic drawing illustrating changes in reward-related activity of SPNs (blue), HFNs (red), and TANs (green) across training phases.

patterns of any of the VMS populations recorded: 81% of the recorded TANs (94/116) exhibited this dynamic signaling across training (Figure S4B).

The detailed structure of the TAN responses (Figures 5D–5F) suggested that the modulation of their activity across training phases largely reflected changes in their response to rewarded outcomes. During the auditory acquisition phase, TANs increased their firing rates for reward and paused in the absence of reward (Figure 5D). During overtraining, the increase at reward nearly disappeared, but the pause at reward omission remained (Figure 5E). Then, during the tactile acquisition phase, their postgoal firing rate increases recovered, and their pauses at reward omission on incorrect trials remained (Figure 5F). Thus, the bidirectional pattern was again present after the contingency switch, so that the TAN responses resembled those of the TANs recorded during the original acquisition of the auditory version of the task. These changes in activity of the reward-related TANs were statistically significant (Figures 5G–5H; Dunn-Sidak post hoc test, $p < 0.05$). We did not detect significant changes in the proportion of reward-responsive TANs among all TANs recorded across the different training stages (Figure 5I; chi-square test; $p > 0.05$). Thus, the rate changes were reflective of the full population of reward-responsive TANs that we recorded. Some TANs (33% of total) showed differences in firing rates for right and left turn trials, but these differences were small (Figure S3F).

The reward-responsive TAN population, like the SPN population, maintained a pronounced response to the warning click. Thus, the sharp changes observed in the firing rates of reward-responsive TANs across the training periods were selective for the outcome period of the task (Figures 5D–5F). To test whether the TAN responses were indeed reward dependent, we trained a control rat on the same task but delayed reward delivery by activating the reward pump 200 ms after goal reaching. The TANs recorded in this animal ($n = 4$) paused their firing at goal reaching on both rewarded and unrewarded trials (Figure 5J). Then, when the delayed reward was delivered in correct trials, they increased their firing rates like the TANs recorded in the standard task. In unrewarded trials, they continued to pause for the absence of the reward, as did TANs in the standard task without reward delay. These results suggest that the reward response was not solely a function of goal reaching on successful trials but was dependent on the availability of reward.

DISCUSSION

Evidence that both reinforcement learning and affective states are influenced by the ventral striatum presents a major challenge to the development of a mechanistic understanding of this striatal region and the circuit-level influence that it exerts: how are each of these functions mediated at the cellular level, and how are they linked? Our findings demonstrate that VMS neurons undergo cell-type-specific patterns of plasticity during reward-based learning that could provide distinct representations of reinforcement contingencies and global task structure.

Of the three identified neuronal populations, only the TANs clearly modulated their reward responses according to changes in the specific reinforcement contingencies presented to the animals and to their performance accuracy (Figure 6). The TANs provided unique bidirectional outcome signals during learning, increasing their firing rate at the end of rewarded trials and decreasing their firing at the end of unrewarded trials. Thus, the TANs reported success or failure of each maze run, according to the current contingencies of the task. These bidirectional TAN responses occurred selectively during the acquisition phases of the sequentially presented tasks. Strikingly, this type of response was not observed during the overtraining phase, when correct performance reached asymptotic levels. During this phase, only the negative responses of the TANs remained. This firing pattern suggests that the bidirectional reward-related responses of the TANs could provide a specific learning signal to ventral striatal circuits.

This sensitivity to switches in reinforcement contingencies exhibited by the VMS TANs contrasted sharply with the relative insensitivity of reward signaling by SPNs in the VMS, which over the course of training gradually reduced their firing at reward (Figure 6). Thus, the SPN population seemed not to contribute to rapid updating of behavior after changes in reinforcement contingencies. Their activity suggested, instead, that they could provide reinforcement signals suitable for participating in initial acquisition of the general motivational or structural aspects of the task (such as the sequence, run through the maze to reach reward), important in establishing the initial learning of the task. Reward-sensitive HFNs also changed their aggregate signaling as training progressed, providing a graded signal potentially suitable for gradually inhibiting the reward responses of the VMS output neurons (Figure 6).

These contrasting cell-type-specific dynamics raise the possibility that during learning, signals related to current task contingencies and signals associated with behavioral policies, considered here as general reward-motivated behaviors and exemplified by running through the maze to reach the goal, could be separately represented and controlled within VMS circuits. We suggest that these dual properties relate not only

to the learning-related functions of the VMS, but also to the pivotal role now attributed to the VMS and its cholinergic interneurons in conditions of negative affect, depression, and anhedonia.

Tonically Active Neurons Provide a Dynamic Learning Signal to VMS Circuits

The dynamic reinforcement-related outcome signaling by TANs in the VMS was notably homogeneous. Over 80% of all recorded TANs exhibited the bidirectional response to reward and nonreward during acquisition phases of training, but after learning exhibited primarily decreases in activity to unrewarded outcomes. TANs in the striatum are widely considered to be cholinergic interneurons (Aosaki et al., 1995; Apicella, 2007; Cragg, 2006; Goldberg and Reynolds, 2011; Wilson et al., 1990). However, recent evidence suggests that similar tonic firing can also be characteristic of the so-called LTS interneurons (putative somatostatin/neuropeptide Y/nitric oxide synthase-containing interneurons) (Beatty et al., 2012). Thus, it is possible that the TANs recorded represented heterogeneous subpopulations, including the 20% of TANs that did not exhibit properties consistent with the majority population reported here. Nevertheless, our comparisons of the firing properties of the TANS to the optogenetically confirmed TANs in ChAT-Cre mice suggest that the population of neurons identified as TANs is likely to have a large component of cholinergic interneurons (English et al., 2012; Witten et al., 2010).

Given this consistency with the properties of identified cholinergic interneurons, it was surprising to find that the typical reward-related response of the VMS TANs was a phasic activation at reward and not the typical pause observed in TANs of the dorsal striatum in response to reward and other motivationally relevant stimuli, accompanying prepause and postpause peaks (Aosaki et al., 1995; Morris et al., 2004; Ravel et al., 2003; Schmitzer-Torbert and Redish, 2008). The bidirectional responding of the TANs in the VMS further set these neurons apart from the conventionally identified TAN populations of the dorsal striatum, many of which pause for both rewarding and aversive conditioning stimuli (Apicella, 2007; Blazquez et al., 2002). Notably, signaling through acetylcholine receptors also differs across the dorsal and ventral striatum (Threlfell and Cragg, 2011). It will be of interest to test the cholinergic identity of these TANs in further behavioral work with reporter-labeled cholinergic neurons, and to compare directly dorsal and ventral striatal TANs in the tasks that we employed here.

In our experiments, we did not explicitly introduce aversive outcomes. However, the absence of reward could be considered aversive under some conditions, including those imposed by the T-maze paradigm. Especially relevant here are the findings of Brown et al. (2012) and Tan et al. (2012), who have found that activation of a GABAergic projection from the VTA to the cholinergic interneurons of the VMS produces pauses in the firing of the cholinergic interneurons and improves behavioral discrimination of aversive cues from neutral ones. It is likely that this cell-type-specific VTA-VMS projection contributed to the pause response of the TANs that we observed during unrewarded trials. Notably, TAN pause responses to reward omission in unrewarded trials were present across all training phases. Thus, the TANs conformed to the model of cholinergic VMS neurons receiving negative reward-related input, but they also exhibited bidirectional signaling of trial outcome during learning, a signal that could strongly drive the learning process in the VMS.

It is possible that these TANs signaled trial completion, rather than trial outcome (correct rewarded or incorrect nonrewarded) per se, but this alternative faces at least two challenges. First, the TANs exhibited bidirectional signaling after the task switch during training stages in which performance was at chance level. Second, this signaling was delayed when the reward was delayed, instead of occurring at trial completion. Yet this possibility cannot be ruled out, especially for the TAN pauses occurring during overtraining when the animals had learned the task. If so, the TAN response could at least at this stage represent a confirmatory response indicative of met expectation (Fujii and Graybiel, 2005).

Learning-Related Neuroplasticity of VMS Spiny Projection Neurons

Given that SPNs are the output neurons of the VMS, it would be natural to expect that they would exhibit strong reward-related activity during acquisition of the initial auditory version of the task, that this activity would subside during overtraining when the behavior became habitual, and that it would reemerge when the new, tactile-cued version of the task was introduced—the pattern that we observed for the majority of TANs. Instead, we found that the SPNs that we identified as exhibiting a selective reward-related response (Apicella et al., 1991; Carelli and Deadwyler, 1994; Taha and Fields, 2005) had a strong reward response during initial acquisition that faded during overtraining and did not react abruptly to the cue switch with a return of reinforcement signaling. The decline observed in reward-related SPN ensemble activity, as well as the decline in the total number of SPNs exhibiting reward-related activity, continued even after the task was switched to the tactile version. By contrast, the warning click and ramp responses of SPNs were unwavering after initial training.

This overall pattern suggests that outcome reporting by the subset of reward-related SPNs declined even as the population of SPNs continued to exhibit responses at trial initiation and goal approach. Thus, once the initial learning had occurred, the SPNs in the VMS largely settled into a pattern that was fairly stable but was lacking the reward-sensitivity that had been apparent during initial training. Such a lack of outcome sensitivity, as demonstrated here physiologically, is a major behavioral feature of developed habits and of addictive states (Balleine and Dickinson, 1998; Belin et al., 2009; Smith et al., 2012).

High-Firing Neurons Include Subpopulations with Opposite Modulation during Learning and Could Serve as a Potential Source of Inhibition of SPN Responses to Reward

We recorded from neurons that fired at high rates, here called HFNs, after the nomenclature of Schmitzer-Torbert and Redish (2008). We were unable to identify these as belonging specifically to either the so-called fast-spiking (FS) or the low-threshold spiking (LTS) GABAergic interneuronal populations that have been described for the dorsal and ventral parts of the striatum

(Berke, 2008; Gittis and Kreitzer, 2012; Kawaguchi et al., 1995; Kreitzer, 2009). These are considered to correspond, respectively, to parvalbumin-containing (FS) interneurons and subsets of somatostatin-containing (LTS) interneurons (Tepper and Bolam, 2004). As we identified subgroups of HFNs by their response properties, the HFNs recorded here could correspond to these or other molecularly distinct subgroups.

The spike patterns of the HFNs that we recorded did clearly place the HFNs as distinct from the SPN and TAN populations that we identified. Over half (58%) of these HFNs were modulated at goal reaching, and these were divided into two different groups. One subgroup of HFNs exhibit increased firing rates on rewarded trials, whereas the second subgroup, comparable to those reported by Lansink et al. (2010), exhibit suppression of their firing rates on rewarded trials. The proportions of neurons in the two subgroups changed inversely and monotonically as training progressed: the percentage of HFNs that paused at reward gradually decreased, whereas the percentage of HFNs that increased their firing during reward gradually increased. Thus, as training progressed, suppression of HFN activity at reward was replaced by increased HFN activity at reward.

These patterns raise the possibility that the HFN subgroups exerted negative control over the gradual changes exhibited by the reward-related SPNs across learning phases. This possibility is consistent with evidence that, in the dorsal striatum, the high-firing neurons identified as FS interneurons exert a feedforward inhibition, contributing to the decreasing reward response observed in SPNs (Kita et al., 1985). It is unlikely, however, that the reward-responsive HFNs controlled all aspects of the SPN activity. For example, the HFNs, unlike the SPNs, did not develop a response to the warning click at trial start. Thus, it is specifically the gradual elevation of HFN responses to reward that could have contributed to the gradual suppression of SPN reward-related activity. Again, repeating these experiments in animals with genetic tags for identification of HFN subtypes would be of great interest.

Could SPNs Represent a Stable Behavioral Policy as a Result of Training?

From the very early stages of training, subsets of SPNs developed phasic increases to the warning click or, as noted by van der Meer and Redish (2011), gradual ramping responses during goal approach. A majority of the ramping neurons also discriminated in their firing rates around the decision period between runs to the left and right end-arms of the maze. Their ramping and click responses, in contrast to their responses at goal reaching, remained relatively stable across training and across the contingency switch, aside from subtle adjustments to the left/right preferences of these cells that we cannot discount. These responses, and most importantly the decline and lack of reactivation of the SPNs at task-version switch, suggest that SPN firing in the VMS during maze runs could be involved in executing a general behavioral policy to initiate movement after the click and to maintain pursuit of reward into specific maze arms. Once established, these global behavioral policies remained relatively static, even in the face of changing reward contingencies, a hallmark of habit formation and drug addiction.

The Influence of VMS Neurons on Habit Learning

Pioneering work on habit learning suggests that the dorsomedial striatum (DMS) is part of a goal-directed learning system that encodes the association of behavioral responses with reward during the early phases of training, when the active seeking of a specific reward leads to encoding of the learned association. Behavioral overtraining, by this view, leads to a shift toward DLS-controlled habitual behavior that can be executed without regard to changes in the identity or value of the reward that originally motivated the learning (Balleine and Dickinson, 1998; Everitt and Robbins, 2005; Thorn et al., 2010; Yin et al., 2005, 2008). We found that reward-related SPN activity in the VMS decreased as training progressed into overtraining but that the SPN population as a whole maintained other responses related to the structure of the task—for example, responses to the warning click and ramping responses. This selective decrease in outcome signaling could contribute, along with decreases in DMS activity, to the emergence of reward-independent stimulus-response habits.

How does the firing of the TANs fit with this view of habit learning? In contrast to the SPNs, the TANs signaled at the end of each trial whether reward was present or not, and this signaling was present during acquisition, fell during overtraining, and reappeared after the cue switch when new learning was required. This pattern suggests that the TANs, putative cholinergic interneurons, can provide an intrinsic learning signal in the VMS. Such a function has also been proposed for TANs in the dorsal striatum of macaques, which can, as a population, faithfully track acquisition, habituation, and reacquisition of a Pavlovian tone-reward contingency (Apicella, 2007; Blazquez et al., 2002). Thus, in both striatal regions, the TANs are poised for a role in modifying local striatal circuits to support the learning of new associations (Aosaki et al., 1994; Bradfield et al., 2013; Centonze et al., 2003; Graybiel et al., 1994; Morris et al., 2004; Pratt and Kelley, 2004; Pratt et al., 2007; Wang et al., 2006). This local trial-by-trial tuning by TANs would be a highly effective teaching signal during the learning of each reinforcement contingency encountered but would not be needed during the phase of habit execution, in which current reward is thought no longer to be the primary driver of behavior. Such modulation according to phase of learning is what we observed.

How this local circuitry is related to shaping the activity of SPNs, however, is still not clear. The fading and reemergence of bidirectional outcome signaling by TANs was not paralleled by similar acquisition-selective responses of reward-related SPNs. This reward-related subpopulation of SPNs exhibited characteristics compatible with modulation by a general familiarity with the reward-related context of the task needed to establish initial goal-directed motivation. It was the HFN population responses and their changing proportions over the course of learning that appeared suited to contribute to the decline in outcome signaling by the SPNs.

Relationship of TAN Firing to Dopaminergic Reward Prediction Error Signals

The increased firing to reward presentation and pause in firing for reward omission observed in TANs are strikingly reminiscent of the patterns of learning-related modulation classically described

for the dopamine-containing neurons of the midbrain (Cohen et al., 2012; Matsumoto and Hikosaka, 2009; Morris et al., 2004; Nakahara et al., 2004; Pan et al., 2005; Roesch et al., 2007; Schultz, 2002; Schultz et al., 1993, 1997). Dopamine-containing neurons have been shown in many studies to track closely mismatches between the predicted outcome and actual outcome (presence or absence of expected reward). In the task employed here, the animals probably relied upon multiple streams of predictive information to determine reward likelihood. Positive reward responses in the TAN population appeared to track reward prediction errors based upon changing performance, indicated by a steady decrease in their responses as performance improved, and a subsequent recovery of responsiveness during relearning. This pattern of signaling could be achieved either directly, by direct access to the reinforcement contingencies, or indirectly, by assessing mismatches between expected and obtained outcome.

HFN and SPN reward-related responses, on the other hand, changed after learning of the first task but failed to recover after the switch in reinforcement contingencies. This failure to recover suggests that reward-related SPNs and HFNs were influenced by predictions based upon unchanging contextual cues or global behavioral policies built up as a result of the acquisition of the first task version (i.e., run to maze end to retrieve reward). This parallel encoding of outcome among different neuronal populations could enable concurrent learning of both reward-associated cues and appropriate reinforcement.

The Activity of Ventral Striatal TANs and Emotion-Related Circuits: A Working Hypothesis

The disruption of VMS cholinergic interneurons has been linked to symptoms of depression, including anhedonia and despair, reflecting a lack of responsiveness to positive or negative outcomes (Warner-Schmidt et al., 2012). Here we demonstrate that VMS TANs have bidirectional outcome responses during habit learning that can encode both positive (reward) and negative (reward omission) outcomes when behavioral change is required. It is striking that the dopamine-containing neurons of the VTA, whose responses are very similar to the TAN responses described here, have also been explicitly implicated in mediating depression-like symptoms in rodents (Tye et al., 2013). Given this evidence, a reasonable possibility is that the VMS TANs, with their bidirectional reinforcement signaling and linkage to the midbrain dopamine neurons, could contribute to coping with new situations (both pleasurable and challenging) and that the absence of this mechanism contributes to the indifference that is a key feature of depression.

Although the bidirectional reinforcement responses of the TANs were selective for the acquisition phases of the tasks, and faded during overtraining, the TANs continued to pause at reward omissions even after long overtraining. The unchanging, noncontingent output signaling by the TAN population thus was in response to negative outcomes. This characteristic of the TAN population suggests that after initial learning, during which the TAN responses represented bivalent signaling, the TANs could contribute to further learning by signaling errors. This property, also, could contribute to a function of these neurons in relation to negative hedonic signaling. If this line of reasoning were extended to normal conditions, the TAN's sensitivity to sudden changes in environmental conditions could be related to emotional reactivity. Such a mechanism favors seeking novelty and new reward, the kind of exploratory behavior that depends on the ventral striatum. Thus, the highly patterned, contingency-sensitive learning signals of the TANs could contribute to aspects of emotional control exerted by ventral striatal circuits.

EXPERIMENTAL PROCEDURES

Long-Evans rats (n = 9) were trained successively on two versions of a T-maze task in which the location of the baited arm was instructed by sensory cues (Figure 1A). In the first version of the task, auditory tones (1 and 8 kHz) were used as cues. After the rats learned (three out four sessions over 72.5% correct performance) and were overtrained (10–15 additional sessions) on the auditory task, they were trained to the same acquisition criterion, followed by 10 additional sessions, on a tactile task in which texture of the maze floor (smooth or rough) signaled reward location. Neural activity in the VMS was recorded from 12 tetrodes (6 in each hemisphere) using a Cheetah data acquisition system (Neuralynx). The position and running speed of the rat in the maze were monitored by an overhead CCD camera and photobeams located throughout the maze. Single units were isolated using Offline Sorter (Plexon) and were classified putatively as SPNs, HFNs, or TANs based on their firing rates and interspike intervals. Spike counts for each unit were then computed in 20 ms bins in a 700 ms window around each task event. These per-unit spike histograms were averaged for each training stage or phase, either with all recorded units in subpopulations examined or after averaging activity of units recorded in multiple sessions (Figures S4C–S4L). Mann-Whitney Wilcoxon rank-sum test, Kruskal-Wallis ANOVA, and Dunn-Sidak post hoc test were used to analyze the changes in neural activity across behavioral training stages with significance set at $p < 0.05$. Recording sites were confirmed in Nissl-stained brain sections. See also Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2014.04.021.

ACKNOWLEDGMENTS

We thank Drs. K.S. Smith, C.A. Thorn, Y. Kubota, V. Lovic, S. Gerstman, M. Platt, and A. Barto for their helpful comments on the manuscript; and H.F. Hall, P. Harlan, C.E. Keller-McGandy, and Dr. D. Hu for their help with the experiments. This work was supported by NIH grant R01 MH060379, by Office of Naval Research grant N00014-04-1-0208, by European Union grant 021716, by CHDI Foundation grant A-5552, by a grant from the Nancy Lurie Marks Family Foundation, by William N. & Bernice E. Bumpus Foundation grant RRDA Pilot: 2013.1, and by Mr. John Wasserlein and Ms. Lucille Braun (A.M.G.).

Accepted: April 11, 2014
Published: June 4, 2014

REFERENCES

Alexander, B., Warner-Schmidt, J., Eriksson, T., Tamminga, C., Arango-Lievano, M., Ghose, S., Vernov, M., Stavarache, M., Musatov, S., Flajolet, M., et al. (2010). Reversal of depressed behaviors in mice by p11 gene therapy in the nucleus accumbens. Sci. Transl. Med. *2*, 54ra76.

Aosaki, T., Graybiel, A.M., and Kimura, M. (1994). Effect of the nigrostriatal dopamine system on acquired neural responses in the striatum of behaving monkeys. Science *265*, 412–415.

Aosaki, T., Kimura, M., and Graybiel, A.M. (1995). Temporal and spatial characteristics of tonically active neurons of the primate's striatum. J. Neurophysiol. *73*, 1234–1252.

Apicella, P. (2007). Leading tonically active neurons of the striatum from reward detection to context recognition. Trends Neurosci. *30*, 299–306.

Apicella, P., Ljungberg, T., Scarnati, E., and Schultz, W. (1991). Responses to reward in monkey dorsal and ventral striatum. Exp. Brain Res. *85*, 491–500.

Atallah, H.E., Lopez-Paniagua, D., Rudy, J.W., and O'Reilly, R.C. (2007). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. Nat. Neurosci. *10*, 126–131.

Balleine, B.W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology *37*, 407–419.

Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature *437*, 1158–1161.

Beatty, J.A., Sullivan, M.A., Morikawa, H., and Wilson, C.J. (2012). Complex autonomous firing patterns of striatal low-threshold spike interneurons. J. Neurophysiol. *108*, 771–781.

Belin, D., and Everitt, B.J. (2008). Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. Neuron *57*, 432–441.

Belin, D., Jonkman, S., Dickinson, A., Robbins, T.W., and Everitt, B.J. (2009). Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. Behav. Brain Res. *199*, 89–102.

Belin, D., Belin-Rauscent, A., Murray, J.E., and Everitt, B.J. (2013). Addiction: failure of control over maladaptive incentive habits. Curr. Opin. Neurobiol. *23*, 564–572.

Berke, J.D. (2008). Uncoordinated firing rate changes of striatal fast-spiking interneurons during behavioral task performance. J. Neurosci. *28*, 10075–10080.

Berridge, K.C., Ho, C.Y., Richard, J.M., and DiFeliceantonio, A.G. (2010). The tempted brain eats: pleasure and desire circuits in obesity and eating disorders. Brain Res. *1350*, 43–64.

Blazquez, P.M., Fujii, N., Kojima, J., and Graybiel, A.M. (2002). A network representation of response probability in the striatum. Neuron *33*, 973–982.

Bradfield, L.A., Bertran-Gonzalez, J., Chieng, B., and Balleine, B.W. (2013). The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. Neuron *79*, 153–166.

Bromberg-Martin, E.S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. Neuron *68*, 815–834.

Brown, M.T., Tan, K.R., O'Connor, E.C., Nikonenko, I., Muller, D., and Lüscher, C. (2012). Ventral tegmental area GABA projections pause accumbal cholinergic interneurons to enhance associative learning. Nature *492*, 452–456.

Carelli, R.M., and Deadwyler, S.A. (1994). A comparison of nucleus accumbens neuronal firing patterns during cocaine self-administration and water reinforcement in rats. J. Neurosci. *14*, 7735–7746.

Centonze, D., Gubellini, P., Pisani, A., Bernardi, G., and Calabresi, P. (2003). Dopamine, acetylcholine and nitric oxide systems interact to induce corticostriatal synaptic plasticity. Rev. Neurosci. *14*, 207–216.

Chang, S.E., Wheeler, D.S., and Holland, P.C. (2012). Roles of nucleus accumbens and basolateral amygdala in autoshaped lever pressing. Neurobiol. Learn. Mem. *97*, 441–451.

Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature *482*, 85–88.

Cragg, S.J. (2006). Meaningful silences: how dopamine listens to the ACh pause. Trends Neurosci. *29*, 125–131.

Der-Avakian, A., and Markou, A. (2012). The neurobiology of anhedonia and other reward-related deficits. Trends Neurosci. *35*, 68–77.

Dezfouli, A., and Balleine, B.W. (2012). Habits, action sequences and reinforcement learning. Eur. J. Neurosci. *35*, 1036–1051.

English, D.F., Ibanez-Sandoval, O., Stark, E., Tecuapetla, F., Buzsáki, G., Deisseroth, K., Tepper, J.M., and Koos, T. (2012). GABAergic circuits mediate the reinforcement-related signals of striatal cholinergic interneurons. Nat. Neurosci. *15*, 123–130.

Everitt, B.J., and Robbins, T.W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat. Neurosci. *8*, 1481–1489.

Fujii, N., and Graybiel, A.M. (2005). Time-varying covariance of neural activities recorded in striatum and frontal cortex as monkeys perform sequential-saccade tasks. Proc. Natl. Acad. Sci. USA *102*, 9032–9037.

Gill, T.M., Castaneda, P.J., and Janak, P.H. (2010). Dissociable roles of the medial prefrontal cortex and nucleus accumbens core in goal-directed actions for differential reward magnitude. Cereb. Cortex *20*, 2884–2899.

Gittis, A.H., and Kreitzer, A.C. (2012). Striatal microcircuitry and movement disorders. Trends Neurosci. *35*, 557–564.

Gläscher, J., Daw, N., Dayan, P., and O'Doherty, J.P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron *66*, 585–595.

Goldberg, J.A., and Reynolds, J.N. (2011). Spontaneous firing and evoked pauses in the tonically active cholinergic interneurons of the striatum. Neuroscience *198*, 27–43.

Gong, S., Doughty, M., Harbaugh, C.R., Cummins, A., Hatten, M.E., Heintz, N., and Gerfen, C.R. (2007). Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. J. Neurosci. *27*, 9817–9823.

Graybiel, A.M. (2008). Habits, rituals, and the evaluative brain. Annu. Rev. Neurosci. *31*, 359–387.

Graybiel, A.M., Aosaki, T., Flaherty, A.W., and Kimura, M. (1994). The basal ganglia and adaptive motor control. Science *265*, 1826–1831.

Greening, S.G., Osuch, E.A., Williamson, P.C., and Mitchell, D.G. (2013). The neural correlates of regulating positive and negative emotions in medication-free major depression. Soc. Cogn. Affect. Neurosci. http://dx.doi.org/10.1093/scan/nst027.

Hernandez, P.J., Sadeghian, K., and Kelley, A.E. (2002). Early consolidation of instrumental learning requires protein synthesis in the nucleus accumbens. Nat. Neurosci. *5*, 1327–1331.

Humphries, M.D., and Prescott, T.J. (2010). The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. Prog. Neurobiol. *90*, 385–417.

Hyman, S.E., Malenka, R.C., and Nestler, E.J. (2006). Neural mechanisms of addiction: the role of reward-related learning and memory. Annu. Rev. Neurosci. *29*, 565–598.

Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V., and Graybiel, A.M. (1999). Building neural representations of habits. Science *286*, 1745–1749.

Kawaguchi, Y., Wilson, C.J., Augood, S.J., and Emson, P.C. (1995). Striatal interneurones: chemical, physiological and morphological characterization. Trends Neurosci. *18*, 527–535.

Keedwell, P.A., Andrew, C., Williams, S.C., Brammer, M.J., and Phillips, M.L. (2005). The neural correlates of anhedonia in major depressive disorder. Biol. Psychiatry *58*, 843–853.

Kita, H., Kita, T., and Kitai, S.T. (1985). Regenerative potentials in rat neostriatal neurons in an in vitro slice preparation. Exp. Brain Res. *60*, 63–70.

Kreitzer, A.C. (2009). Physiology and pharmacology of striatal neurons. Annu. Rev. Neurosci. *32*, 127–147.

Kreitzer, A.C., and Malenka, R.C. (2008). Striatal plasticity and basal ganglia circuit function. Neuron *60*, 543–554.

Kubota, Y., Liu, J., Hu, D., DeCoteau, W.E., Eden, U.T., Smith, A.C., and Graybiel, A.M. (2009). Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum. J. Neurophysiol. *102*, 2142–2160.

Lansink, C.S., Goltstein, P.M., Lankelma, J.V., and Pennartz, C.M. (2010). Fast-spiking interneurons of the rat ventral striatum: temporal coordination

of activity with principal cells and responsiveness to reward. Eur. J. Neurosci. *32*, 494–508.

Li, J., and Daw, N.D. (2011). Signals in human striatum are appropriate for policy update rather than value prediction. J. Neurosci. *31*, 5504–5511.

Liljeholm, M., and O'Doherty, J.P. (2012). Contributions of the striatum to learning, motivation, and performance: an associative account. Trends Cogn. Sci. *16*, 467–475.

Matsumoto, M., and Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature *459*, 837–841.

McClure, S.M., Berns, G.S., and Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. Neuron *38*, 339–346.

Morris, G., Arkadir, D., Nevet, A., Vaadia, E., and Bergman, H. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. Neuron *43*, 133–143.

Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y., and Hikosaka, O. (2004). Dopamine neurons can represent context-dependent prediction error. Neuron *41*, 269–280.

Nestler, E.J., and Carlezon, W.A., Jr. (2006). The mesolimbic dopamine reward circuit in depression. Biol. Psychiatry *59*, 1151–1159.

Pan, W.X., Schmidt, R., Wickens, J.R., and Hyland, B.I. (2005). Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. J. Neurosci. *25*, 6235–6242.

Pratt, W.E., and Kelley, A.E. (2004). Nucleus accumbens acetylcholine regulates appetitive learning and motivation for food via activation of muscarinic receptors. Behav. Neurosci. *118*, 730–739.

Pratt, W.E., Spencer, R.C., and Kelley, A.E. (2007). Muscarinic receptor antagonism of the nucleus accumbens core causes avoidance to flavor and spatial cues. Behav. Neurosci. *121*, 1215–1223.

Price, J.L., and Drevets, W.C. (2010). Neurocircuitry of mood disorders. Neuropsychopharmacology *35*, 192–216.

Ravel, S., Legallet, E., and Apicella, P. (2003). Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. J. Neurosci. *23*, 8489–8497.

Roesch, M.R., Calu, D.J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat. Neurosci. *10*, 1615–1624.

Schmitzer-Torbert, N.C., and Redish, A.D. (2008). Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype. Neuroscience *153*, 349–360.

Schultz, W. (2002). Getting formal with dopamine and reward. Neuron *36*, 241–263.

Schultz, W., Apicella, P., Ljungberg, T., Romo, R., and Scarnati, E. (1993). Reward-related activity in the monkey striatum and substantia nigra. Prog. Brain Res. *99*, 227–235.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.

Setlow, B., Schoenbaum, G., and Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning. Neuron *38*, 625–636.

Shabel, S.J., Proulx, C.D., Trias, A., Murphy, R.T., and Malinow, R. (2012). Input to the lateral habenula from the basal ganglia is excitatory, aversive, and suppressed by serotonin. Neuron *74*, 475–481.

Smith, K.S., Virkud, A., Deisseroth, K., and Graybiel, A.M. (2012). Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. Proc. Natl. Acad. Sci. USA *109*, 18932–18937.

Stamatakis, A.M., and Stuber, G.D. (2012). Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. Nat. Neurosci. *15*, 1105–1107.

Svenningsson, P., Chergui, K., Rachleff, I., Flajolet, M., Zhang, X., El Yacoubi, M., Vaugeois, J.M., Nomikos, G.G., and Greengard, P. (2006). Alterations in 5-HT1B receptor function by p11 in depression-like states. Science *311*, 77–80.

Taha, S.A., and Fields, H.L. (2005). Encoding of palatability and appetitive behaviors by distinct neuronal populations in the nucleus accumbens. J. Neurosci. *25*, 1193–1202.

Tan, K.R., Yvon, C., Turiault, M., Mirzabekov, J.J., Doehner, J., Labouèbe, G., Deisseroth, K., Tye, K.M., and Lüscher, C. (2012). GABA neurons of the VTA drive conditioned place aversion. Neuron *73*, 1173–1183.

Tepper, J.M., and Bolam, J.P. (2004). Functional diversity and specificity of neostriatal interneurons. Curr. Opin. Neurobiol. *14*, 685–692.

Thorn, C.A., Atallah, H., Howe, M., and Graybiel, A.M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. Neuron *66*, 781–795.

Threlfell, S., and Cragg, S.J. (2011). Dopamine signaling in dorsal versus ventral striatum: the dynamic role of cholinergic interneurons. Front. Syst. Neurosci. *5*, 11.

Tobler, P.N., Dickinson, A., and Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. J. Neurosci. *23*, 10402–10410.

Tye, K.M., Mirzabekov, J.J., Warden, M.R., Ferenczi, E.A., Tsai, H.C., Finkelstein, J., Kim, S.Y., Adhikari, A., Thompson, K.R., Andalman, A.S., et al. (2013). Dopamine neurons modulate neural encoding and expression of depression-related behaviour. Nature *493*, 537–541.

van der Meer, M.A., and Redish, A.D. (2011). Ventral striatum: a critical look at models of learning and evaluation. Curr. Opin. Neurobiol. *21*, 387–392.

van der Meer, M.A., Johnson, A., Schmitzer-Torbert, N.C., and Redish, A.D. (2010). Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. Neuron *67*, 25–32.

Voorn, P., Vanderschuren, L.J., Groenewegen, H.J., Robbins, T.W., and Pennartz, C.M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. Trends Neurosci. *27*, 468–474.

Wacker, J., Dillon, D.G., and Pizzagalli, D.A. (2009). The role of the nucleus accumbens and rostral anterior cingulate cortex in anhedonia: integration of resting EEG, fMRI, and volumetric techniques. Neuroimage *46*, 327–337.

Wang, Z., Kai, L., Day, M., Ronesi, J., Yin, H.H., Ding, J., Tkatch, T., Lovinger, D.M., and Surmeier, D.J. (2006). Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. Neuron *50*, 443–452.

Warner-Schmidt, J.L., Schmidt, E.F., Marshall, J.J., Rubin, A.J., Arango-Lievano, M., Kaplitt, M.G., Ibañez-Tallon, I., Heintz, N., and Greengard, P. (2012). Cholinergic interneurons in the nucleus accumbens regulate depression-like behavior. Proc. Natl. Acad. Sci. USA *109*, 11360–11365.

Wickens, J.R., Horvitz, J.C., Costa, R.M., and Killcross, S. (2007). Dopaminergic mechanisms in actions and habits. J. Neurosci. *27*, 8181–8183.

Wilson, C.J., Chang, H.T., and Kitai, S.T. (1990). Firing patterns and synaptic potentials of identified giant aspiny interneurons in the rat neostriatum. J. Neurosci. *10*, 508–519.

Witten, I.B., Lin, S.C., Brodsky, M., Prakash, R., Diester, I., Anikeeva, P., Gradinaru, V., Ramakrishnan, C., and Deisseroth, K. (2010). Cholinergic interneurons control local circuit activity and cocaine conditioning. Science *330*, 1677–1681.

Yin, H.H., and Knowlton, B.J. (2006). The role of the basal ganglia in habit formation. Nat. Rev. Neurosci. *7*, 464–476.

Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005). The role of the dorsomedial striatum in instrumental conditioning. Eur. J. Neurosci. *22*, 513–523.

Yin, H.H., Ostlund, S.B., and Balleine, B.W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. Eur. J. Neurosci. *28*, 1437–1448.